*Themed Issue: Process Analytical Technology*
*Guest Editor - Ajaz Hussain*

# Process Analytical Technology Case Study: Part II. Development and Validation of Quantitative Near-Infrared Calibrations in Support of a Process Analytical Technology Application for Real-Time Release

*Submitted: December 13, 2004; Accepted: April 19, 2005; Published: October 6, 2005*

Robert P. Cogdill,[1] Carl A. Anderson,[1] Miriam Delgado,[1] Robert Chisholm,[2] Raymond Bolton,[2] Thorsten Herkert,[3] Ali M. Afnan,[4] and James K. Drennen III[1]

[1]Duquesne University Center for Pharmaceutical Technology, Pittsburgh, PA 16066
[2]AstraZeneca, Macclesfield, Cheshire, SK10 4TF, UK
[3]AstraZeneca GmbH, Plankstadt, 68723, Germany
[4]US Food and Drug Administration, Center for Drug Evaluation and Research, Office of Pharmaceutical Science, HFD-003, Rockville, MD 20852
*The views presented in this article do not necessarily reflect those of the Food and Drug Administration.

## ABSTRACT

This article is the second of a series of articles detailing the development of near-infrared (NIR) methods for solid dosage-form analysis. Experiments were conducted at the Duquesne University Center for Pharmaceutical Technology to demonstrate a method for developing and validating NIR models for the analysis of active pharmaceutical ingredient (API) content and hardness of a solid dosage form. Robustness and cross-validation testing were used to optimize the API content and hardness models. For the API content calibration, the optimal model was determined as multiplicative scatter correction with Savitsky-Golay first-derivative preprocessing followed by partial least-squares (PLS) regression including 4 latent variables. API content calibration achieved root mean squared error (RMSE) and root mean square error of cross validation (RMSECV) of 1.48 and 1.80 mg, respectively. PLS regression and baseline-fit calibration models were compared for the prediction of tablet hardness. Based on robustness testing, PLS regression was selected for the final hardness model, with RMSE and RMSECV of 8.1 and 8.8 N, respectively. Validation testing indicated that API content and hardness of production-scale tablets is predicted with root mean square error of prediction of 1.04 mg and 8.5 N, respectively. Explicit robustness testing for high-flux noise and wavelength uncertainty demonstrated the robustness of the API concentration calibration model with respect to normal instrument operating conditions.

**Corresponding Author:** James K. Drennen, School of Pharmacy, Duquesne University, Pittsburgh, PA 15282. Tel: (412) 396-5520; Fax: (412) 396-5130; E-mail: drennen@duq.edu

**KEYWORDS:** PAT, process analytical technology, near-infrared spectroscopy, chemometrics, pharmaceutical analysis

## INTRODUCTION

Process analytical technology (PAT) promises to provide benefits to both manufacturers and consumers of pharmaceutical products. Such technology will improve manufacturing efficiency, while enhancing process understanding. The advances in pharmaceutical manufacturing that are possible from enhanced process understanding will augment the quality of pharmaceuticals, while reducing the cost of manufacturing, and, ultimately, will play an important role in controlling the price of health care.[1,2]

This article is the second of a series of articles detailing the development of near-infrared (NIR) methods for solid dosage form analysis from feasibility studies through implementation. The first article described efforts to qualify the capabilities of instrumentation and sample handling systems, evaluate the potential effect of process signature on calibration development, and compare the utility of reflection and transmission data collection methods. This article describes the development and validation of quantitative NIR calibrations for active ingredient content and hardness, including the selection and optimization of chemometric techniques. The third article will follow with the philosophy and methodology of implementing and managing a PAT application to ensure the continuity of performance, including a detailed description of performance monitoring and calibration transfer techniques for the analytical method. The experimental results to be covered in this series demonstrate that, for this application, a properly developed NIR method provides predictive analyses with comparable accuracy and precision to the reference methods currently used, while reducing the time required for analysis.

NIR spectroscopy is an inherently multivariate analytical technique. The broad, overlapping NIR absorbance bands contain redundant chemical and physical information that can be used for the prediction of many product quality attributes from a single NIR spectrum. Specialized signal processing and multivariate statistical data analysis techniques, known as chemometrics, are essential to the implementation of many process analytical technologies. Chemometrics are often used for multivariate statistical process control[3,4] and exploratory data analysis. These techniques often rely on latent variable factor analysis (eg, principal component analysis) and are described in the literature.[5-7] Inverse regression techniques, such as multilinear regression, principal component regression,[5,8] and partial least-squares (PLS)[9] are used for the development of prediction models relating product quality attributes to spectral data.

The on-line analytical method developed and validated in this article is intended to be integrated into the release testing of a well-established product. NIR scans of selected tablets will be acquired on-line, after compression and before the coating of tablets. Spectral data from manufactured tablets are to be processed using the predictive models described in this article. Results will be used to assess the real-time state of control of the manufacturing process and product quality attributes. Active pharmaceutical ingredient (API) content and hardness of tablets will be continuously predicted by the NIR method during manufacturing. The product studied for this investigation has been manufactured for several years and has a history of nominal variability in measured product quality attributes.[10] It is anticipated that this analytical method will be integrated with a series of other analytical methods and controls as a PAT application for real-time release.

Because the calibrations were developed for a well-established product, as opposed to a product in formulation, a wealth of production samples were available for method development. The product is manufactured to very tight specifications, providing insufficient variability to create a stable calibration model. To remedy this, production sample sets were augmented with laboratory-scale sample sets[10] to enhance variability in API content and hardness. Samples produced in the laboratory were designed to not only enhance the variability of the product quality attributes for this calibration but were similar enough to production samples to allow model development based on NIR data from manufacturing and laboratory samples.[10]

For many pharmaceutical solid dosage forms, NIR spectroscopy is an ideal PAT tool for assessing critical product quality parameters. NIR spectroscopy is used to nondestructively analyze tablets by measuring their transmission or reflection spectra[11,12] at many wavelengths within the NIR region of the electromagnetic spectrum. Using modern, high-speed instrumentation, noncontact analyses are performed in <1 s, with a high degree of accuracy and precision.[13-17]

The current analytical methods for this solid dosage form include off-line high-performance liquid chromatography (HPLC) and diametral crushing strength tests. The NIR methods are intended to replace the current analytical tests. Each NIR calibration must meet individual requirements for accuracy, precision, linearity, specificity, and robustness. Because the NIR instrumentation, tablet samples, and the spectral data acquisition procedure used in this application were described in an earlier article,[10] this article is focused on quantitative method development and validation.

### *Objective*

Although the combination of NIR spectroscopy and chemometrics is well established, their use in pharmaceutical quality assurance brings new opportunities and challenges. Building on the work presented in the first article in this series,[10] the objective of this article is to demonstrate a method for developing and validating NIR models for the analysis of API content and hardness of a solid dosage form.

## MATERIALS AND METHODS

The method was developed for the analysis of 50-mg (nominal) tablet cores before finishing and release to market. All of the calculations were performed using Matlab 6.5 (The Mathworks Inc, Natick, MA), along with the PLS_Toolbox 3.0 (Eigenvector Research Inc, Wenatchee, WA) and other analysis tools written at the Duquesne University Center for Pharmaceutical Technology for this study.

### *API Model Development*

Model development for API content used 500 samples drawn from 10 production and 13 laboratory-scale batches. A full-factorial experimental design of 10 batches and a fractional factorial design of 13 batches were manufactured in production-scale and laboratory-scale settings, respectively. In both designs, the factors varied were as follows: source of API (3 vendors), source of magnesium stearate (3 vendors), moisture content at granulation (2 levels), and compression force (2 levels). Additionally, for the laboratory-scale experiments, the API content was varied over 9 levels from 70% to 130% (of nominal API content). Reference data for both the production and laboratory-scale data were obtained using reverse-phase HPLC. The SE of the laboratory method has been observed to be approximately 3% of nominal API content (approximately 1.5 mg). The calibration datasets are summarized in Table 1. The API

**Table 1.** Summary of API Content Calibration and Validation Statistics

| | Calibration Dataset | VAL1 | VAL2 | VAL3 |
|---|---|---|---|---|
| Samples (n) | 500 | 350 | 40 | 38 |
| Batches (n) | 23 | 30 | 4 | 2 |
| Maximum (mg) | 65.32 | 66.06 | 50.31 | 65.53 |
| Mean (mg) | 48.90 | 49.07 | 49.44 | 47.92 |
| Minimum (mg) | 32.66 | 33.63 | 47.70 | 32.43 |
| Standard Deviation (mg) | 5.83 | 4.94 | 0.67 | 13.93 |
| Model Type | | Full-specturm PLS regression | | |
| Preprocessing | | MSC + 1st Derivative | | |
| Spectral Range (nm) | | (1300 – 2000), 2 | | |
| Latent Variables (n) | | 4 | | |
| RMSE (mg)* | 1.48 | 1.25 | 5.35 (1.04) | 5.07 (3.76) |
| RMSE (%, nominal)* | 2.96 | 2.50 | 10.7 (2.08) | 10.1 (7.52) |
| r | 0.967 | 0.972 | 0.441 | 0.974 |
| $r^2$ | 0.936 | 0.944 | 0.194 | 0.948 |
| RPD* | 3.9 | 4.0 | NA | 2.7 (3.8) |
| Bias (mg) | 0.00 | −0.22 | −5.3 (0.71) | −4.0 (2.04) |

*A prediction bias was identified for the VAL2 and VAL3 datasets. The corrected values are in parentheses.

content validation data consisted of three subsets selected from 43 laboratory-scale and 39 production-scale lots.

The first validation set (VAL1) consisted of 350 samples; 5 batches of laboratory-scale samples were included to provide extended range, along with 25 production batches. The production batches were manufactured at a different facility (facility A) than the API content calibration samples (facility B). The second validation set (VAL2) consisted of 40 production samples manufactured at facility A. Furthermore, the spectra were collected at a much later date, after a teardown and reassembly of the instrument. The third validation (VAL3) set consisted of 38 laboratory-scale samples manufactured at extreme API content levels, with significant variation in the force of compression. The purpose of the third validation set is to additionally test the linearity and robustness of the calibration over a wide range of physical and chemical variations. The API content validation datasets are summarized in Table 1.

All of the calibration spectra were collected in the laboratory using a Brimrose, Luminar 3070 (Brimrose Corp, Baltimore, MD) scanning acoustooptic tunable filter spectrometer described in the prior work.[10] Measurements were conducted on tablets at room temperature (approximately 20°C). The absorbance spectra were truncated to remove the noisy regions of the spectrum and retain important spectral features. The 700-nm window between 1,300 and 2,000 nm was selected, with a 2-nm wavelength interval (Figures 1 and 2). The truncated absorbance spectra are shown in Figure 3.

The spectral preprocessing steps used for the API content calibration were selected based on simulated robustness testing and cross-validation. Both the API content and hardness calibrations were developed using PLS regression. PLS regression was chosen (as opposed to multilinear regression or principal component regression) because it is a proven, industry-standard factor regression method, which is easily optimized using readily available statistical analysis tools. Furthermore, as with other factor analysis methods, PLS regression analysis generates a set of orthogonal basis vectors, or "loadings," which provide a useful platform for model interpretation and outlier detection. Batch-wise cross-validation was performed to select the appropriate number of factors (latent variables) to be

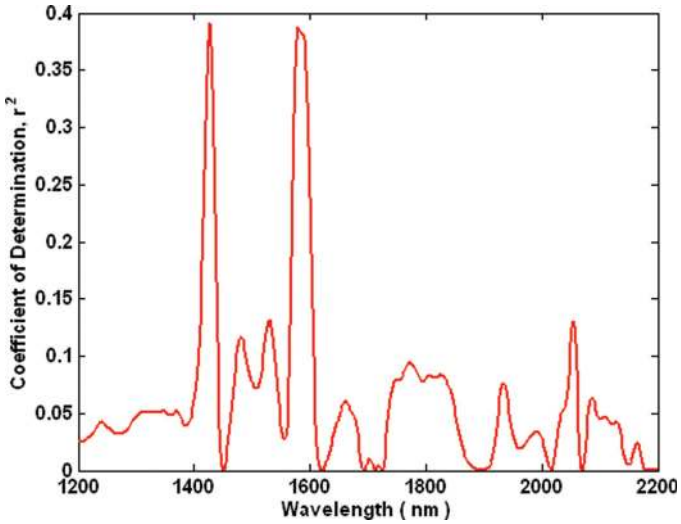


**Figure 1.** Coefficient of determination between NIR absorbance and API content as a function of wavelength. The relatively low coefficient of determination at the ends of the spectrum suggest the wavelength range can be truncated.

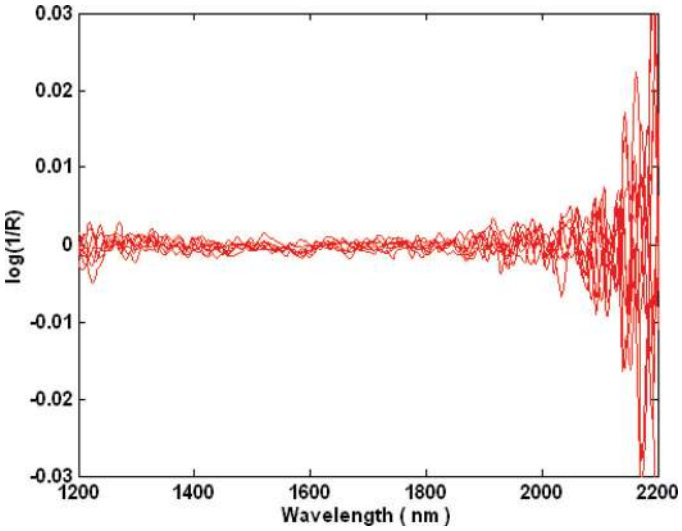


**Figure 2.** High-flux spectrometer noise as a function of wavelength, measured by repeatedly scanning a reflectance standard (without repositioning). Based on the rapid increase in noise toward the ends of the spectrum, the data were truncated at 1,300 and 2,000 nm.

included in the PLS model. For batch-wise cross-validation, the calibration set was divided into a number of subsets where each subset (23 in total) contained samples from a single lot of tablets.

### *Hardness*

The range of hardness for the calibration dataset was varied by changing the compression force across 3 levels for 15 laboratory-scale batches and 7 production-scale lots (Table 2). Hardness reference data were collected by diametral crushing-strength test of the tablet cores. The diametral crushing-strength test was performed using model

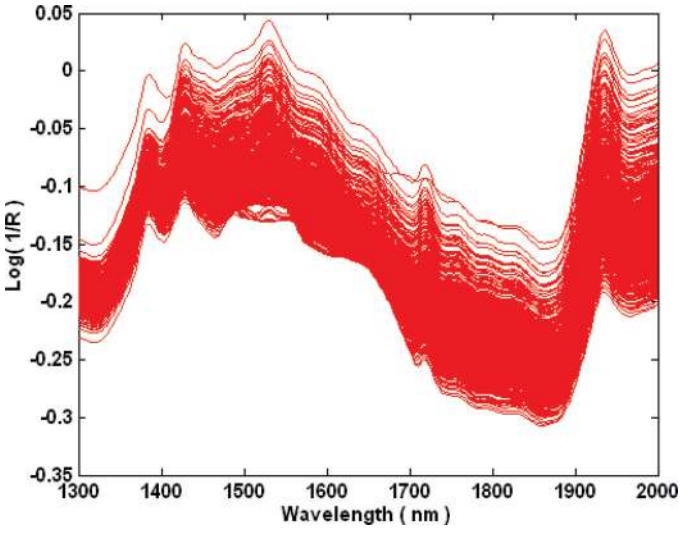


**Figure 3.** Truncated NIR absorbance spectra from the API content calibration dataset.

**Table 2.** Summary of Hardness Calibration and Validation Dataset Statistics

| | Calibration Dataset | Validation Dataset |
|---|---|---|
| Samples (n) | 437 | 152 |
| Batches (n) | 22 | 8 |
| Maximum (N) | 140.0 | 145.0 |
| Mean (N) | 61.7 | 58.1 |
| Minimum (N) | 16.0 | 13.0 |
| Standard Deviation (N) | 29.2 | 30.9 |
| Model Type | Full-specturm PLS regression | |
| Preprocessing | MSC+1$^{st}$ Derivative | |
| Spectral Range (nm) | (1300 – 2000), 2 | |
| Latent Variables (n) | 3 | |
| RMSE (N)* | 8.1 | 12.0 (8.5) |
| r | 0.961 | 0.961 |
| $r^2$ | 0.922 | 0.92344 |
| RPD* | 3.6 | 2.6 (3.6) |
| Bias (N) | 0.0 | −8.0 (−0.01) |

*A prediction bias was identified. Corrected values are in parentheses.

UTS-12F hardness tester, manufactured by Charles Ischi, AG (Zuchwil, Germany). The hardness validation data set was comprised of 5 laboratory-scale and 3 production-scale lots.

The hardness calibration model was optimized in the same manner as the API content models. Preprocessing routines were selected before batch-wise cross-validation, wherein the number of latent variables included in the model was determined. For comparison, a second hardness calibration method was tested using a spectral best fit algorithm based on fitting the baseline, slope, and curvature of the absorbance spectra.[17] Baseline-fit models using first-order and second-order baseline parameters were tested.

## RESULTS AND DISCUSSION

### *Method Development*

The first article in this series evaluated several fundamental elements of method development for on-line measurements in PAT applications. Those elements included the poolability of laboratory-scale and production-scale data, the impact of the process sampling method, and the choice between reflection and transmission measurements. The poolability of samples from various scales of manufacturing is a critical issue in method development. Samples from multiple production scales are used to provide a sufficiently broad range in critical product parameters for robust calibration development. Appropriate process sampling is an essential element of an on-line method. For this application, tablet positioning was the primary concern of

process sampling. The decision between reflection and transmission is important for NIR-based methods. This article focuses on development and validation of reflection-based NIR methods for API content and tablet hardness analyses.

### *API Content*

The choice of spectral preprocessing is an important part of the NIR method optimization process. Appropriate preprocessing may reduce the impact of sources of variability, such as instrumental drift and sampling variability, correct for pathlength variation, and increase the signal-to-background level. Appropriate selection of signal processing facilitates model accuracy and robustness.[18] Many NIR applications are successfully optimized with a few basic operations, typically baseline/scatter correction and smoothing derivatives.[6,7,12] Because the tablet spectra would be acquired in the diffuse reflectance mode, variation in the spectral baseline was expected because of the effect of variation in the sample-to-analyzer geometry on the scatter portion of the signal, as well as a result of the nonlinear pathlength effects associated with diffuse reflectance. Thus, the optimal preprocessing routine was expected to include a scatter correction using the standard-normal-variate or multiplicative scatter correction (MSC) transformations. Smoothing derivatives were tested in combination with and without scatter correction as a means of amplifying the portions of the spectrum with relevant variation. A quantitative approach was undertaken to help select the most suitable combination of spectral preprocessing operations for the API calibration. This approach considers accuracy and robustness in the calculation of a robustness index (RI) as described below.

Several approaches for calculating RIs have been proposed.[18] The RI used for this work is defined as the inverse of the area under the curve defined by a quadratic fit of root mean squared error (RMSE) plotted as a function of additional spectral noise during robustness testing (Figure 4).

$$\mathrm{RI} = \frac{1}{\int_{L_N=0}^{3} \mathrm{A}L_N^2 + \mathrm{B}L_N + \mathrm{C}} \qquad (1)$$

where $L_N$ is the level of simulated noise added and $\mathrm{A}L_N^2 + \mathrm{B}L_N + \mathrm{C}$ is the quadratic fit of the noise-augmented prediction error data. This equation assesses the error of the model across a range of simulated noise levels. To perform the test, the prediction accuracy of a PLS calibration was tested in a Monte Carlo fashion at numerous levels of additional noise (from 0 to 3 SDs). RI was calculated from the increase in prediction error across simulated noise levels using Equation 1. The best preprocessing combinations were identified as those with the highest RI (and lowest root mean square error of cross validation (RMSECV)). Noise was generated using randomly weighted combinations of noise factors with additional Gaussian "white" noise. The noise factors were generated by principal components analysis of difference spectra between repeat scans, scans of the same sample over a long period of time, and the difference spectra between scans of the same sample on multiple instruments. In this way, robustness to sources of variability, such as sampling error, and instrument drift can be evaluated. The results of the robustness testing for the preprocessing combinations tested are shown in Table 3. A combination of MSC and Savitsky-Golay 1st Derivative was chosen based on the maximum RI, which was concurrent with prior experience with these data. Because the RI method is being used only as a screening tool, no effort has been made to calculate levels of significant difference between treatments. The mean spectrum of the calibration set was used as the reference for application of MSC; a 19-point, second order polynomial smoothing filter was applied during the smoothing-derivative operation. The preprocessed calibration spectra are shown in Figure 5.

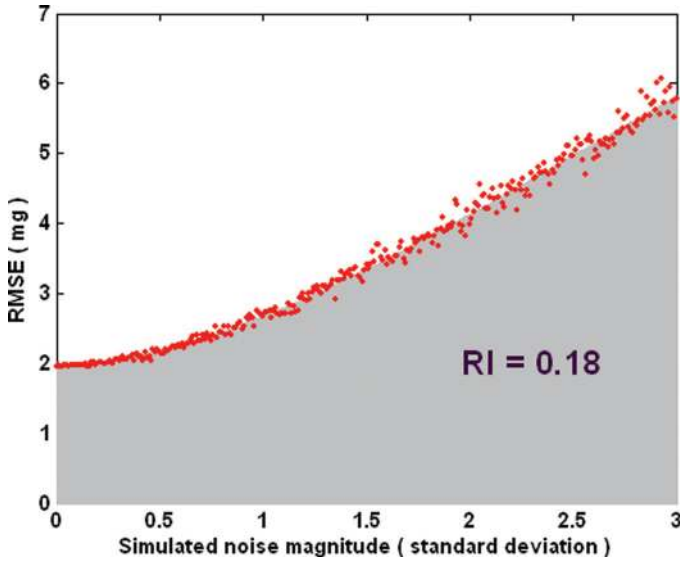


**Figure 4.** RI test result for raw (unpreprocessed) NIR spectra and API content calibration.

A batch-wise cross-validation study was performed. The ideal number of PLS factors to be included in the final model is indicated by a minimum RMSECV or the maximum number of PLS factors before the decrease in the RMSECV is negligible. Batch-wise cross-validation increases confidence that the model will perform well when applied to new samples. It was determined that the API content calibration should include 4 PLS factors (Figures 6 and 7).

Important spectral features in the PLS factor loadings were identified to provide evidence for model specificity

**Table 3.** Results of Cross-Validation and RI Testing of Preprocessing Combinations Tested for the API Content Calibration*

| Preprocessing Treatment | PLS Factors | RMSECV (mg) | RMSE (mg) | $r^2$ | Robutness Index |
|---|---|---|---|---|---|
| Raw Data | 5 | 2.36 | 1.96 | 0.886 | 0.18 |
| SNV | 5 | 2.19 | 1.70 | 0.915 | 0.25 |
| SNV + 1st Deriv. | 4 | 1.79 | 1.49 | 0.935 | 0.33 |
| SNV + 2nd Deriv. | 3 | 1.93 | 1.64 | 0.921 | 0.29 |
| MSC | 5 | 2.12 | 1.68 | 0.917 | 0.26 |
| MSC + 1st Deriv. | 4 | 1.80 | 1.48 | 0.936 | 0.33 |
| MSC + 2nd Deriv. | 3 | 1.89 | 1.61 | 0.923 | 0.29 |
| 1st Deriv. | 4 | 1.80 | 1.48 | 0.936 | 0.32 |
| 2nd Deriv. | 3 | 1.89 | 1.61 | 0.924 | 0.29 |

*A Larger RI Score Indicates a More Robust Calibration. Only the Relative Magnitudes of the RI Scores Are of Value, Because Other Applications Will Have Incompatible Scales.

(Figure 8). The significant features in the regression vector (Figure 9) are attributed to API and lactose absorbance.[10] It should be noted that because the calibration was calculated using first-derivative spectra, the zero-crossings of the regression vector correspond to the center of major absorbance bands.

For this study, API content is predicted as the mass of API in each tablet. A more rigorous approach to the development of a calibration for API content would be based on concentration of API. However, the actual weight of each tablet used for calibration was not available. Rather than assuming nominal weight for each tablet, the API content calibration was derived in terms of API mass (milligrams). Using the nominal mass would add an error to the calibration that is equivalent to the normal variation in tablet mass.

### *Hardness*

For this work, 2 empirical methods were compared for development of the hardness calibration: PLS regression and baseline-fit. Whereas concentration calibrations (eg, API content and moisture content) rely on a linear response between absorbance and concentration, the measurement of physical parameters, such as tablet hardness, is a less direct analysis. A number of earlier works have addressed the issue of measuring tablet hardness via NIR spectroscopy.[15,17,19-21] It was found in these studies that variation in tablet hardness is manifested in NIR spectra as a nonlinear baseline shift that increases with wavelength. The wavelength dependency is attributed to the multiple-scatter effect that causes higher baseline absorbance at longer wavelengths. This effect is commonly observed in diffuse reflectance NIR spectroscopy of particulate solids. This

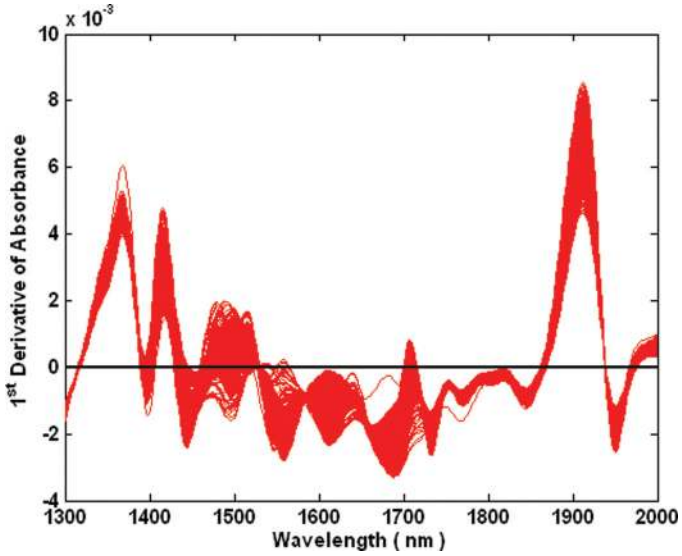


**Figure 5.** Calibration spectra after MSC and Savitsky-Golay first-derivative preprocessing. The derivative was calculated using a second-order polynomial fit over a 36-nm (19-point) window.

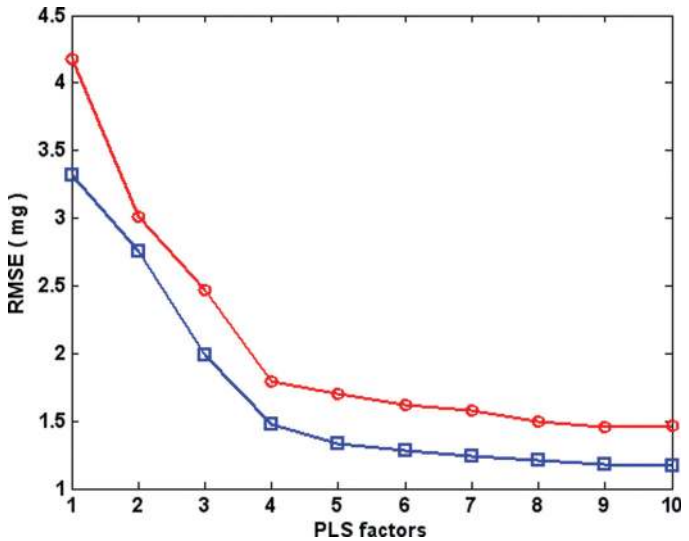


**Figure 6.** Batch-wise cross-validation for optimization of the API content PLS calibration. The upper curve (circles) corresponds to cross-validation error; the lower curve (squares) corresponds to RMSE of calibration.

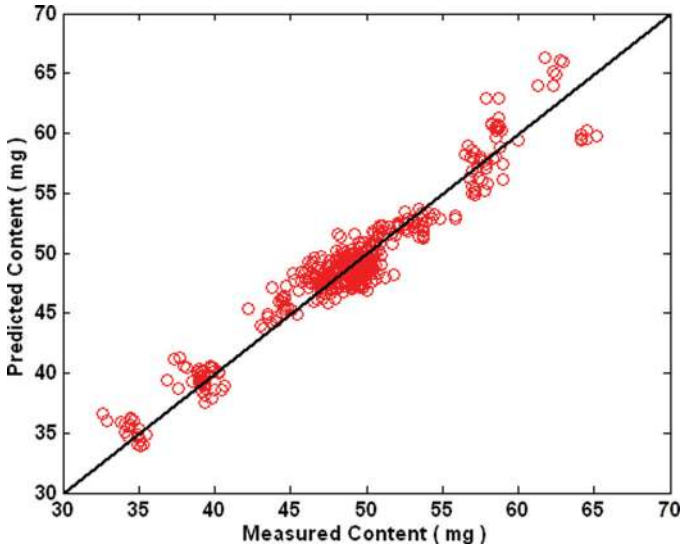


**Figure 7.** Plot of actual versus predicted API content from the API content calibration data set.

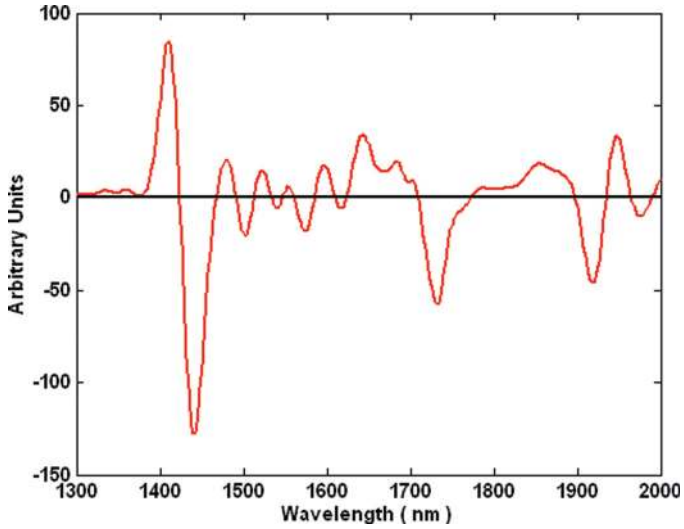


**Figure 9.** API content calibration regression coefficient vector. Because the calibration spectra were transformed to first derivative, the major zero-crossings of the vector indicate the center of significant absorbance bands. The zero-crossings near 1,430 nm is the center of the major API absorbance band.

suggests some form of a Kubelka-Munk scattering relationship, similar to that observed with particle size variation.[22,23]

A baseline-fit method for the prediction of hardness was developed. This method of calibration development requires a different process than PLS. Rather than correlating tablet hardness to NIR absorbance at each wavelength using multivariate linear regression, a polynomial function is fitted through each NIR spectrum, and the hardness reference data are then correlated to the coefficients of the polynomial equations. Prior work[17] has focused on fitting a first-order polynomial and then either correlating hardness to the slope coefficient or using multivariate regression to correlate hardness to both coefficients. For this work, 4 different baseline-fit calibrations were derived. Two first-order calibrations were created: one used both the zero-order and first-order coefficients, and the other correlated hardness to only the first-order coefficient. Two second-order baseline calibrations were created: one correlated all 3 of the polynomial coefficients to hardness, and the other used only the first-order and second-order baseline fit coefficients. Rather than fitting a polynomial to the entire spectrum for the baseline-fit calibrations, a significant improvement in accuracy was observed by selecting only a portion of the spectrum for fitting. The center wavelength and width of the spectral window selected for each calibration was chosen by minimizing RMSE using an exhaustive search algorithm.

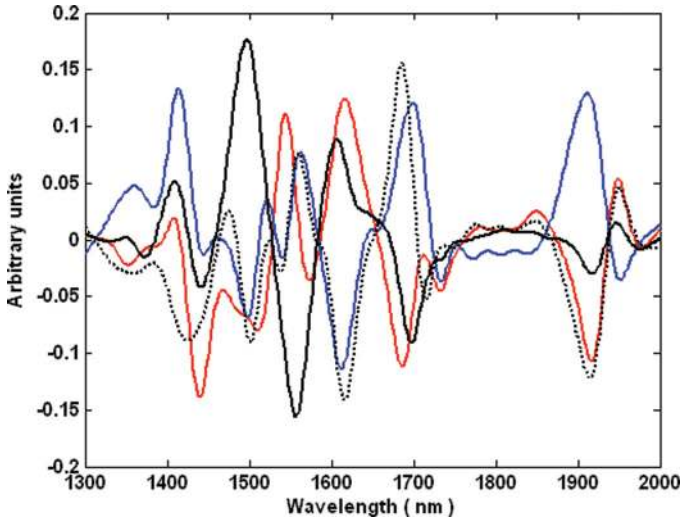


**Figure 8.** PLS factor loadings for the API content calibration (red = factor 1, blue = factor 2, black = factor 3, black dashed = factor 4).

In the same manner as the API content calibration, the final hardness calibration was developed by optimizing the RI (Table 4). Whereas the baseline fit calibrations performed similarly to the PLS calibrations in terms of calibration error (RMSE), the PLS regression calibrations were found to be superior in terms of RI. The combination of MSC and first derivative was selected for its balance of accuracy and robustness. Although both the API content and hardness calibrations used the same preprocessing treatment, no correlation between their respective predictions was observed when provided the same spectra. The regression vector for the hardness calibration is shown in Figure 10. Interpretation of the spectral features in the regression vector is difficult for the hardness calibrations (Figure 11). It is important to remember that PLS is a form of pattern recognition. With appropriate validation and parallel testing, the calibration can be deployed as an effective alternative to destructive testing.

**Table 4.** Results of cross-validation and robustness index (RI) testing of preprocessing combinations and model types considered for the hardness calibration. A larger RI score indicates a more robust calibration

| Preprocessing Treatment | Model Type | Factors/ Terms | RMSECV | RMSE | $r^2$ | Robustness Index |
|---|---|---|---|---|---|---|
| Raw Data | PLS | 3 | 10.11 | 8.88 | 0.907 | 0.042 |
| SNV | PLS | 3 | 10.96 | 9.75 | 0.888 | 0.040 |
| SNV 1st Deriv. | PLS | 3 | 12.04 | 10.86 | 0.861 | 0.038 |
| SNV 2nd Deriv. | PLS | 3 | 12.38 | 11.08 | 0.855 | 0.037 |
| MSC | PLS | 3 | 9.66 | 8.82 | 0.908 | 0.043 |
| MSC 1st Deriv. | PLS | 3 | 8.82 | 8.11 | 0.923 | 0.047 |
| MSC 2nd Deriv. | PLS | 3 | 8.21 | 7.48 | 0.934 | 0.046 |
| 1st Deriv. | PLS | 2 | 9.11 | 8.22 | 0.920 | 0.043 |
| 2nd Deriv. | PLS | 3 | 8.73 | 7.91 | 0.926 | 0.045 |
| 1st Order | Baseline Fit | 2 | NA | 8.79 | 0.909 | 0.036 |
| 1st Order | Baseline Fit | 1 | NA | 8.75 | 0.910 | 0.038 |
| 2nd Order | Baseline Fit | 3 | NA | 8.01 | 0.924 | 0.033 |
| 2nd Order | Baseline Fit | 2 | NA | 8.32 | 0.918 | 0.038 |

### *Method Validation*

After the specification of analytical method parameters and calculation of calibration equations, the efforts of method development must shift away from experimentation and optimization toward validation. Through the use of independent testing and statistical data analysis, the validation of NIR spectroscopic methods must determine the following 5 key performance criteria: accuracy, specificity, linearity, precision, and robustness. Whereas the API content and hardness calibrations are validated separately, both use the same validation procedure. The specificity of the calibrations is addressed in the method development sections of this article.

### *API Content*

Accuracy and linearity of the API calibration were assessed by statistical analysis of API predictions. A summary of API content validation statistics is shown in Table 1. During validation testing, it was observed that predictions from VAL2 and VAL3 were significantly biased. Given the history of stability that the method had shown, it was surmised that the bias was a result of the intercontinental transport and reassembly of the NIR analyzer. Instrumental drift because of ageing (years) contributes to the bias. Because the instrumental changes occurred after the calibration and VAL1 data were collected, only VAL2 and VAL3 were affected. The bias was calculated using an independent set of 8 production-scale tablets. A bias correction of 6 mg was calculated. The bias adjustment applied to VAL2 and VAL3 is applied to all of the subsequent data analyses.

After correction of the instrument bias, satisfactory method performance was observed for all 3 of the validation datasets (Table 1). A RMSE of prediction <1.5 mg

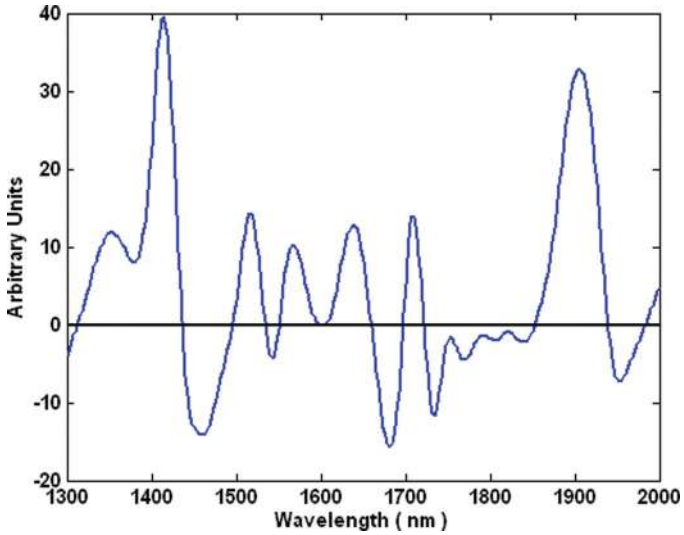


**Figure 10.** Hardness calibration regression coefficient vector.

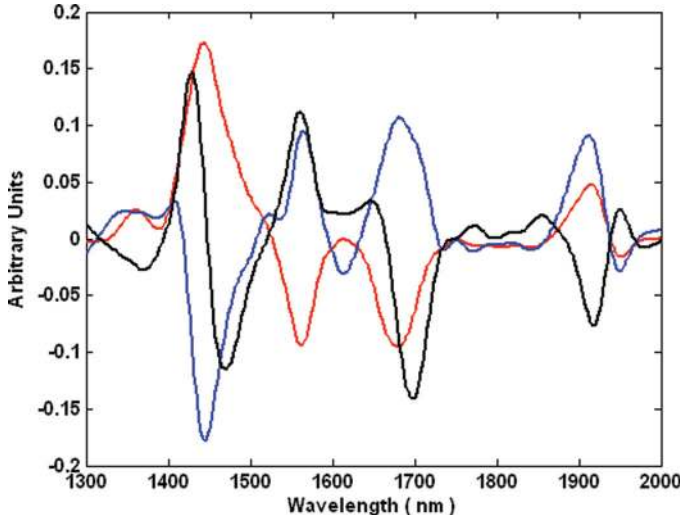


**Figure 11.** PLS factor loadings for the hardness calibration (red = factor 1, blue = factor 2, black = factor 3).

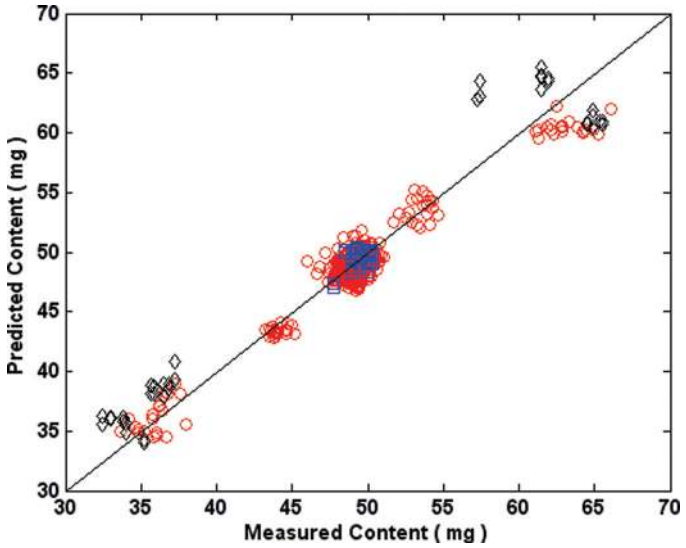


**Figure 12.** Prediction plot for API content validation datasets (VAL1 = circles, VAL2 = squares, VAL3 = diamonds). The VAL2 and VAL3 predictions have been corrected for bias.

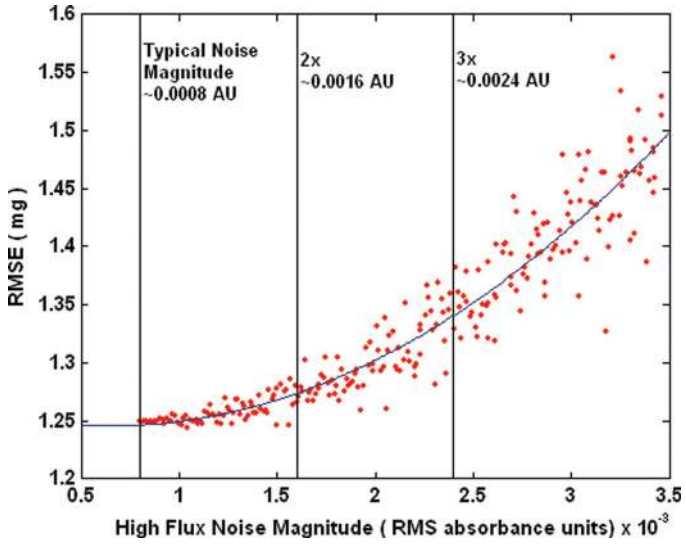


**Figure 13.** High-flux noise robustness test results.

(3% nominal) for VAL1 and VAL2 was calculated. The data used to calculate a coefficient of determination ($r^2$) of 0.93 for VAL1 and VAL3 are displayed in Figure 12. Analysis of the coefficient of determination for VAL2 is irrelevant, because the SD of API at production-scale (0.67 mg)[10] is less than half the SE of the HPLC reference (approximately 3% nominal API content/1.5 mg). Furthermore, the SD of API content for the VAL3 dataset (13.93 mg) exceeds the SD of the calibration dataset (4.94 mg). The relative performance determinant (the ratio of standard error of prediction to SD of modeled API content) is near 4. This suggests that the calibration is suitable for quantitative prediction.[24]

Precision of the API content analysis was assessed by measuring the SD of predicted values obtained by replicate analyses[10] of a set of 16 tablets (n = 160). Method precision was calculated as the mean within-group SD of predicted API content. Using batch-wise cross-validation, confidence limits were calculated as ±1.5 times the SD of precision. The precision of the API content calibration was found to be 0.5 ± 0.015 mg.

Robustness was built into the API content calibration model by including hundreds of samples with considerable diversity (raw material supplier, production line, date of analysis, etc.). Method robustness was implicitly and explicitly demonstrated. Calibration robustness was verified implicitly by the stability of the calibration in predicting validation data with intentional variation in potentially confounding factors. An explicit demonstration of method robustness was performed by measuring the effect of simulated high-flux noise and wavelength axis variation. These studies facilitate development of rational specifications for instrument performance. Both the high-flux noise and wavelength accuracy robustness simulations were performed by adding simulated noise to the VAL1 dataset and measuring the effect on the predictive performance of the API content calibration. High-flux noise was simulated by adding vectors of normally distributed random numbers to the VAL1 spectra. It should be noted that some level of high-flux noise and wavelength inaccuracy is present in the data before the addition of simulated noise. The intrinsic level of high-flux noise in the spectra was found to be approximately 0.0008 absorbance units. The SD of the random noise at each wavelength was scaled to match the true noise profile of the instrument (Figure 2). Wavelength axis error was simulated by shifting the VAL1 spectra to lower and higher frequencies in small increments via cubic spline interpolation. The results of high-flux noise and wavelength accuracy robustness testing are shown in Figures 13 and 14.

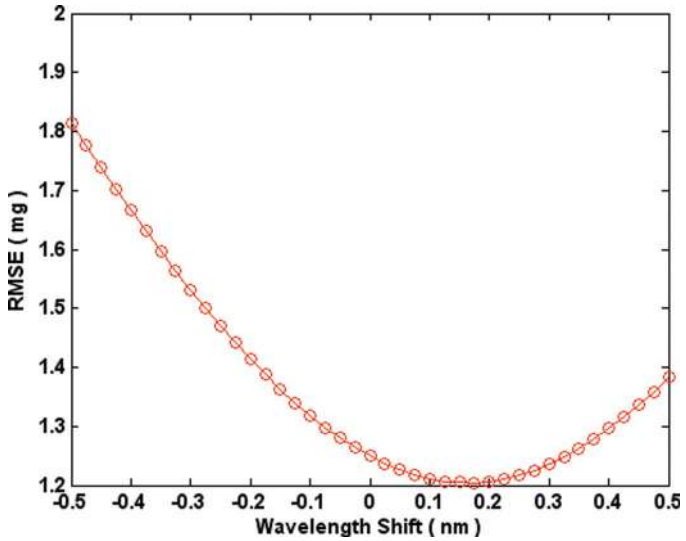


**Figure 14.** Wavelength accuracy robustness test results.

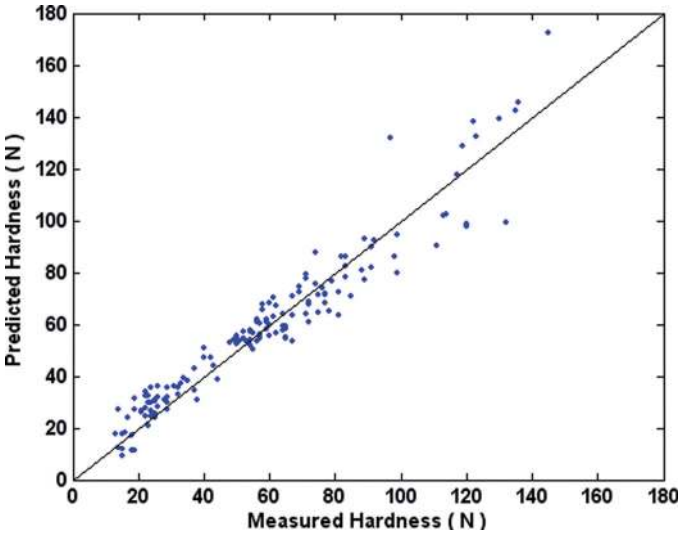


**Figure 15.** Hardness validation prediction plot (predicted values have been corrected for bias).

The level of high-flux noise and wavelength shift used in validation testing must exceed the limits of normal operation. Doing so provides assurance that the method will not fail during minor deviations from normal instrumental operating conditions. In the event that instrument performance degrades significantly from the limits of normal operation, instrument failure should be detected by continuous calibration monitoring.

### *Hardness*

The hardness validation data and validation test results for the hardness calibration are summarized in Table 2 and Figure 15. As was observed during validation of the API content calibration, a bias correction was needed for the hardness validation data, which were collected after the instrumental variations described above. The data collected for validation testing was used to calculate a bias correction. Before the hardness calibration can be deployed, a subsequent independent dataset for validation testing is required.

The precision of the hardness calibration, using the same dataset and method used for precision testing (10 replicates of 16 samples) of the API content calibration, was estimated to be 9.6 N for hardness predictions. This is similar in magnitude to the RMSE calculated for the hardness predictions of 8.5 N. Historically, the crushing-strength test has a precision of approximately 9 N. The similarity in the precision, error, and reference methods indicates that the error of the reference method is responsible for the precision and error of the NIR prediction. A summary of the hardness calibration dataset statistics is shown in Table 2.

## CONCLUSION

The objective of this article was to demonstrate a method for developing and validating NIR models for the analysis of API content and hardness of a solid dosage form. Models for the prediction of API content and hardness were developed and optimized. Calibration model form (for hardness) and preprocessing operations were selected based on RI analysis and cross-validation testing. Independent datasets were used to demonstrate method accuracy, precision, linearity, specificity, and robustness. Tests were performed to explicitly demonstrate API calibration robustness to variation in instrumental high-flux noise and wavelength shift. The next installment in the series describes the procedures for calibration monitoring and calibration transfer.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of Ken Leiper (Benson Associates, Caythorpe, UK) for his valuable effort in the conception of the work and experimental design used in this and subsequent articles. Additional data collection and analysis were performed by David Molseed. This work was funded through an agreement between Duquesne University Center for Pharmaceutical Technology and AstraZeneca.

Tables 1 to 4 and Figures 7 and 13 are reprinted with permission from Advanstar Communications.[25]

## REFERENCES

1. US Food and Drug Administration. PAT: A Framework for Innovative Manufacturing and Quality Assurance, Guidance for Industry. US Food and Drug Administration, Washington, DC: 2003.

2. US Food and Drug Administration. Guidance Documents: Website of FDA Guidances. US Food and Drug Administration: 2004.

3. Box GEP, Jenkins GM, Reinsel GC. *Time Series Analysis.* Upper Saddle River. NJ: Prentice Hall; 1994.

4. Jackson JE, Mudholkar GS. Control procedures for residuals associated with principal components analysis. *Technometrics.* 1979;21:341-349.

5. Martens H, Naes T. *Multivariate Calibration.* New York, NY: John Wiley and Sons; 1989.

6. Mobley PR, Kowalski BR, Workman JJ Jr, Bro R. Review of chemometrics applied to spectroscopy: 1985-95, Part 2. *Appl Spectrosc Rev.* 1996;31:347-368.

7. Workman JJ Jr, Mobley PR, Kowalski BR, Bro R. Review of chemometrics applied to spectroscopy: 1985-95, Part I. *Appl Spectrosc Rev.* 1996;31:73-124.

8. Naes T, Martens H. *Principal component regression in NIR analysis*: viewpoints, background details, and selection of components. *J Chemometric.* 1988;2:155-167.

9. Wold S, Ruhe A, Wold H, Dunn WJ III. The collinearity problem in linear regression. the partial least squares approach to generalized inverses. *SIAM J Sci Stat Comput.* 1984;5:735-743.

10. Cogdill RP, Anderson CA, Delgado-Lopez M, et al. Process analytical technology case study: Part I. feasibility studies for quantitative NIR method development. *AAPS PharmSciTech.* 2005;E263-273.

11. Ingle JD, Crouch SR. *Spectrochemical Analysis.* Englewood Cliffs, NJ: Prentice-Hall; 1988.

12. Williams P, Norris K, eds. *Near-Infrared Technology in the Agricultural and Food Industries.* 2nd ed. St Paul, MN: American Association of Cereal Chemists; 2001.

13. Drennen JK, Lodder RA. Nondestructive near-infrared analysis of intact tablets for determination of degradation products. *J Pharm Sci.* 1990;79:622-627.

14. Iyer M, Morris H, Drennen JK. Solid dosage form analysis by near infrared spectroscopy: Comparison of reflectance and transmittance measurements including the determination of effective sample mass. *J Near Infrared Spectrosc.* 2002;10:233-245.

15. Kirsch JD, Drennen JK. Determination of film-coated tablet parameters by near-infrared spectroscopy. *J Pharm Biomed Anal.* 1995;13:1273-1281.

16. Kirsch JD, Drennen JK. Near-infrared spectroscopy: applications in the analysis of tablets and solid pharmaceutical dosage forms. *Appl Spectrosc Rev.* 1995;30:139-174.

17. Kirsch JD, Drennen JK. Nondestructive tablet hardness testing by near-infrared spectroscopy: A new and robust spectral best-fit algorithm. *J Pharm Biomed Anal.* 1999;19:351-362.

18. Zeaiter M, Roger JM, Bellon-Maurel V, Rutledge DN. Robustness of models developed by multivariate calibration. Part I: The assessment of robustness. *Trends Anal Chem.* 2004;23:157-170.

19. Bijlani VJ, Delgado-Lopez M, Adeyeye CM, Drennen JK. Near infrared spectroscopy for monitoring the hardness of roller compaction ribbons. *NIR News.* 2002;13:8-14.

20. Donoso M, Kildsig DO, Ghaly ES. Prediction of tablet hardness and porosity using near-infrared diffuse reflectance spectroscopy as a nondestructive method. *Pharm Dev Technol.* 2003;8:357-366.

21. Drennen JK, Voytilla RJ. Non-destructive tablet hardness testing: the effect of moisture on tablet hardness prediction. *NIR News.* 2002;13:12-13.

22. Bull CR. Compensation for particle size effects in near-infrared reflectance. *Analyst.* 1991;116:781-786.

23. Norris KH, Williams PC. Optimization of mathematical treatments of raw near-infrared signals in the measurement of protein in hard red spring wheat, I: Influence of Particle Size. *Cereal Chem.* 1984;61:158-165.

24. AACC. Near-Infrared Methods: Guidlines for Model Development and Maintenance. In: *Approved Methods of the American Association of Cereal Chemists.* Washington, DC: AACC Press; 1999; AACC Method 39-00.

25. Cogdill RP, Anderson CA, Drennen JK III. Integrated PAT Method Development. In: *Process Analytical Technology, a supplement to Pharmaceutical Technology.* Iselin, NJ: Advanstar Communications; 2004;29-34.